



Validation numérique de l'approche d'analyse de sensibilité spatiale de Lilburne et Tarantola

Nathalie Saint-Geours

► To cite this version:

Nathalie Saint-Geours. Validation numérique de l'approche d'analyse de sensibilité spatiale de Lilburne et Tarantola. 2010. hal-00570318

HAL Id: hal-00570318

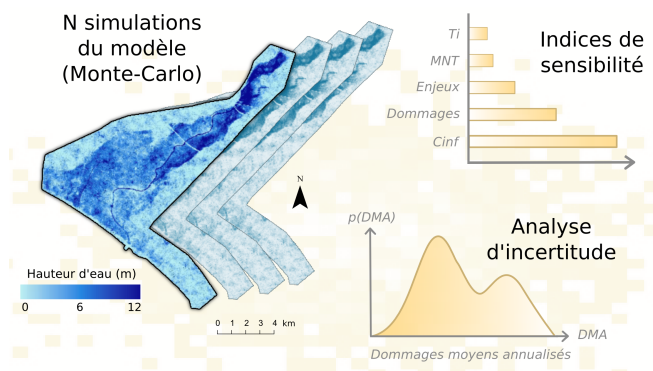
<https://hal.science/hal-00570318>

Submitted on 28 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Validation numérique de l'approche d'analyse de sensibilité spatiale de Lilburne et Tarantola



Résultats

- 28 février 2011 -

Nathalie Saint-Geours

Sommaire

1	Fonctions tests	1
1.1	Champ spatial aléatoire $Z(\theta, \phi)$	1
1.2	Fonctions tests.	3
2	Estimation des indices de sensibilité	4
2.1	Méthode L	4
2.2	Méthodes RD et RDG	4
2.3	Méthode S	5
3	Résultats	6
3.1	Comparaison des quatres méthodes d'estimation.	6
3.2	Convergence de l'approche de Lilburne & Tarantola	7
4	Discussion	7
4.1	Comparaison des quatres méthodes d'estimation.	7
4.2	Convergence des estimateurs de Lilburne avec n .	9
5	Conclusion	9
	Bibliographie	22
	Annexes.	23

Introduction

On s'intéresse ici aux diverses méthodes proposées pour évaluer la sensibilité d'une sortie de modèle $Y = f(X_1, \dots, X_k)$ à l'incertitude qui pèse sur un facteur d'entrée X_i distribué spatialement. On se restreint aux approches sans construction d'une surface de réponse (méta-modélisation). L'objet de ce document est :

- de valider de manière empirique la méthode proposée par Lilburne & Tarantola [3]
- de comparer (de manière numérique) cette méthode avec les autres méthodes

On étudie pour cela trois fonctions tests. Le document est structuré de la manière suivante : la section 1 présente les fonctions tests ; la section 2 décrit les méthodes d'estimation des indices de Sobol ; la section 3 présente les résultats numériques obtenus ; la section 4 discute en détails ces résultats.

1 Fonctions tests

On considère trois fonctions tests, que l'on nomme f_1 , f_2 et f_3 . Ces trois fonctions prennent en entrée deux facteurs scalaires X_1 et X_2 , et un facteur d'entrée distribué spatialement, noté Z . La loi de probabilité retenue pour X_1 et X_2 varie selon la fonction test considérée. En revanche, le modèle d'incertitude retenu pour le facteur spatial Z est le même pour les trois fonctions tests.

1.1 ■ Champ spatial aléatoire $Z(\theta, \phi)$

On considère en entrée des trois fonctions test étudiées un même champ spatial aléatoire Z défini par :

$$\begin{aligned} Z : [-90, 90]^2 &\rightarrow \mathbb{R} \\ (\theta, \phi) &\mapsto Z(\theta, \phi) \end{aligned} \tag{1}$$

Z est un champ spatial 2D aléatoire généré par la fonction `campbell12D`. Cette fonction est issue d'un papier de Campbell et al. (2006), elle a ensuite été modifiée par B. Iooss et A. Marrel (Marrel et al., 2010). Elle prend en entrée huit facteurs d'entrée scalaires C_1 à C_8 et renvoie en sortie un champ 2D indexé par des coordonnées polaires θ et ϕ . Les paramètres C_1 à C_8 sont des variables aléatoires réelles supposées indépendantes et de loi uniforme dans $[-1, 5]$. La fonction `campbell12D` a été conçue pour donner un champ avec de fortes hétérogénéités spatiales, et une structure très différente selon les valeurs des paramètres C_i (voir fig.1). On a :

$$\begin{aligned} Z(\theta, \phi) &= g(C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, \theta, \phi) \\ &= C_1 \exp - \frac{(0.8\theta + 0.2\phi - 10C_2)^2}{60C_1^2} + (C_2 + C_4) \exp \frac{(0.5\theta + 0.5\phi)C_1}{500} \\ &\quad + C_5(C_3 - 2) \exp \frac{-(0.4\theta + 0.6\phi - 20C_6)^2}{40C_5^2} + (C_6 + C_8) \exp \frac{(0.3\theta + 0.7\phi)C_7}{250} \end{aligned} \tag{2}$$

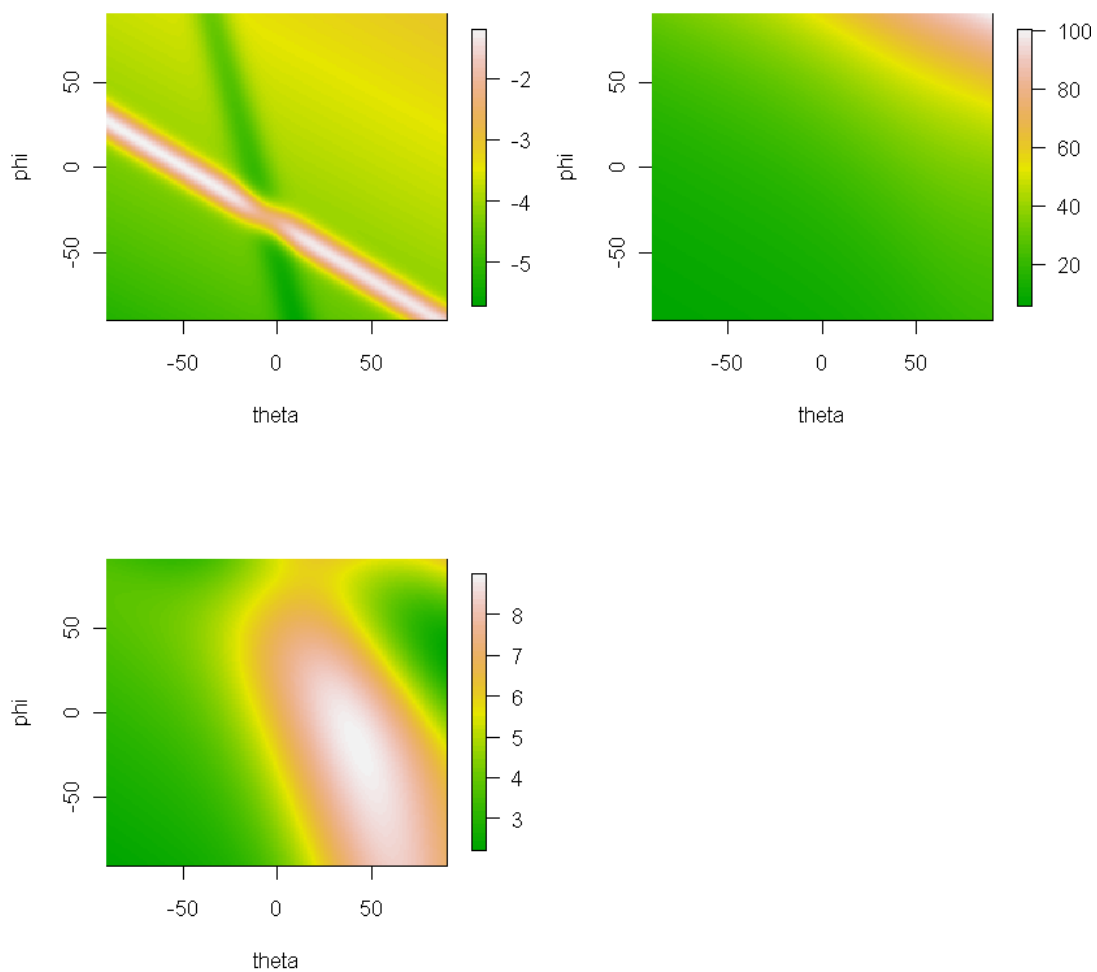


FIG. 1 – Trois cartes Z générées par la fonction Campbell2D :
 $C = (-1, -1, -1, -1, -1, -1, -1, -1)$ (left) ; $C = (5, 5, 5, 5, 5, 5, 5, 5)$ (center) ; $C = (5, 3, 1, -1, 5, 3, 1, -1)$ (right).

1.2 ■ Fonctions tests

Les trois fonctions test f_1 , f_2 et f_3 considérées prennent trois facteurs d'entrée : X_1 , X_2 (qui sont tous deux scalaires), et Z (champ 2D). Ces fonctions sont des «*point-based models*», ou encore des modèles «*verticaux*» : la valeur de la sortie Y est la somme sur l'ensemble du domaine spatial $(\theta, \phi) \in [-90, 90]^2$ d'une même fonction f évaluée en chaque point ; la valeur de la fonction f au point (θ, ϕ) ne dépend que des facteurs scalaires X_1 , X_2 et de la valeur du champ Z au point (θ, ϕ) .

Fonction test f_1

La sortie Y_1 est la somme sur l'ensemble du domaine $(\theta, \phi) \in [-90, 90]^2$ de la fonction d'Ishigami f_1 estimée en chaque point (θ, ϕ) avec pour facteurs d'entrée X_1 , X_2 et $Z(\theta, \phi)$:

$$\begin{aligned} Y_1 &= \sum_{(\theta, \phi) \in [-90, 90]^2} f_1(X_1, X_2, Z(\theta, \phi)) \\ &= \sum_{(\theta, \phi) \in [-90, 90]^2} \sin[Z(\theta, \phi)] + A \cdot \sin(X_1)^2 + B \cdot X_2^4 \cdot \sin[Z(\theta, \phi)] \end{aligned} \quad (3)$$

où :

$$\begin{cases} A = 7 \\ B = 0.1 \\ X_1 \sim \mathcal{U}[-\pi, \pi] \\ X_2 \sim \mathcal{U}[-\pi, \pi] \end{cases} \quad (4)$$

Fonction test f_2

La sortie Y_2 est la somme sur l'ensemble du domaine $(\theta, \phi) \in [-90, 90]^2$ de la fonction affine f_2 estimée en chaque point (θ, ϕ) avec pour facteurs d'entrée X_1 , X_2 et $Z(\theta, \phi)$:

$$\begin{aligned} Y_2 &= \sum_{(\theta, \phi) \in [-90, 90]^2} f_2(X_1, X_2, Z(\theta, \phi)) \\ &= \sum_{(\theta, \phi) \in [-90, 90]^2} X_1 \cdot Z(\theta, \phi) + X_2 \end{aligned} \quad (5)$$

où

$$\begin{cases} X_1 \sim \mathcal{U}[0.5, 1.5] \\ X_2 \sim \mathcal{U}[-8, 8] \end{cases} \quad (6)$$

Fonction test f_3

La sortie Y_3 est la somme sur l'ensemble du domaine $(\theta, \phi) \in [-90, 90]^2$ de la fonction f_3 (fonction G de Sobol) estimée en chaque point (θ, ϕ) avec pour facteurs d'entrée X_1 , X_2 et $Z(\theta, \phi)$:

$$\begin{aligned} Y_3 &= \sum_{(\theta, \phi) \in [-90, 90]^2} f_3(X_1, X_2, Z(\theta, \phi)) \\ &= \sum_{(\theta, \phi) \in [-90, 90]^2} \left(\frac{|4Z(\theta, \phi) - 2| + a_0}{1 + a_0} \cdot \prod_{i=1}^2 \frac{|4X_i - 2| + a_i}{1 + a_i} \right) \end{aligned} \quad (7)$$

où

$$\begin{cases} a_0 = 0 \\ a_1 = 1 \\ a_2 = 4.5 \\ X_1 \sim \mathcal{U}[-5, 5] \\ X_2 \sim \mathcal{U}[-0.15, 0.15] \end{cases} \quad (8)$$

2 Estimation des indices de sensibilité

On compare quatre méthodes d'estimation des indices de sensibilité, que l'on nomme L (méthode proposée par Lilburne & Tarantola [3]), RD (réduction de la dimension), RDG (réduction de la dimension et groupe) et S (méthode usant d'une variable «switch»). On les décrit brièvement ici.

2.1 ■ Méthode L

La méthode L est celle proposée par Lilburne & Tarantola [3]. On suit la procédure suivante :

1. on génère en amont de l'analyse de sensibilité un jeu de n réalisations aléatoires de la carte d'entrée Z . Pour cela, on tire un échantillon de taille n contenant chacun des paramètres C_1 à C_8 tirés dans leur loi de probabilité, puis on évalue la fonction `Campbell2D` pour chaque ligne i de cet échantillon, générant ainsi une réalisation $Z^{(i)}$ du champ Z .
2. les n réalisations $Z^{(i)}$ de la carte Z sont considérées équiprobables, et chacune est associée à un unique entier i compris entre 1 et n .
3. on tire ensuite deux échantillons A et B de même taille M mêlant les facteurs d'entrée scalaires X_1 , X_2 et une variable aléatoire discrète \tilde{Z} qui prend pour valeurs, de manière équiprobable, les entiers compris entre 1 et n . Pour chaque ligne de l'échantillon, la valeur de \tilde{Z} indique le numéro de réalisation aléatoire de la carte Z à utiliser pour évaluer la fonction test.
4. on estime les indices de Sobol de X_1 , X_2 et de \tilde{Z} par la méthode de Saltelli et al. [4] à partir des deux échantillons A et B .

Le nombre d'évaluation de la fonction test nécessaire à l'estimation des indices est dans ce cas :

$$N = 2 \cdot (3 + 1) \cdot M = 8 \cdot M$$

Par ailleurs, le nombre d'appels à la fonction `Campbell2D` qui génère le champ Z est ici égal à n .

2.2 ■ Méthodes RD et RDG

Les méthodes RD et RDG s'appuient sur une réduction de la dimension du champ Z . L'idée est de résumer l'information contenue dans le facteur d'entrée spatialisé Z en un nombre restreint de facteurs scalaires que l'on peut considérer indépendants. Ici cette réduction de la dimension est artificielle, puisqu'il nous suffit de considérer les huit facteurs d'entrée C_1 à C_8 de la fonction `Campbell2D` qui permet de générer le champ Z . Les C_i sont supposés indépendants entre eux.

Pour quantifier la sensibilité de la sortie Y à l'incertitude dans le champ Z , deux approches peuvent être utilisées : on peut estimer les indices de sensibilité (de premier ordre et totaux) de chacun des facteurs scalaires C_i ; ou bien on peut calculer un indice de sensibilité de premier ordre et un indice total pour le groupe de facteurs (C_1, \dots, C_8) . On nomme la première approche RD pour «réduction de la dimension», c'est notamment la technique retenue par Volokova et al. [5] ou par Busby et al. [1]. La seconde approche sera appelée RDG , pour «réduction de la dimension et groupe».

Méthode *RD*

Pour la méthode *RD*, on suit la procédure suivante :

1. on tire deux échantillons *A* et *B* de même taille *M* mêlant X_1, X_2 et C_1 à C_8 tirés chacun dans leur loi de probabilité.
2. on estime les indices de Sobol de $X_1, X_2, C_1, C_2, C_3, C_4, C_5, C_6, C_7$ et C_8 par la méthode de Saltelli et al. [4] à partir des deux échantillons *A* et *B* ; pour chaque ligne *i* de l'échantillon, la valeur des paramètres C_1 à C_8 pour cette ligne permet de calculer une carte $Z^{(i)}$ à l'aide de la fonction `Campbell2D`.

Le nombre d'évaluations de la fonction test est :

$$N = 2 \cdot (10 + 1) \cdot M = 22 \cdot M$$

Le nombre d'appels à la fonction `Campbell2D` est *N*.

Méthode *RDG*

La méthode *RDG* est identique à la méthode *RD*, à ceci près qu'elle regroupe l'ensemble des paramètres C_1 à C_8 dans un même «groupe de facteurs». Un indice de Sobol unique est estimé pour l'ensemble du groupe (voir le paragraphe 1.2.15 de Saltelli et al. [4] pour une discussion de l'analyse de Sobol par groupes). On suit la procédure suivante :

1. on tire deux échantillons *A* et *B* de même taille *M* mêlant X_1, X_2 et C_1 à C_8
2. on estime les indices de Sobol de X_1, X_2 , et un indice unique pour le groupe ($C_1, C_2, C_3, C_4, C_5, C_6, C_7$ et C_8) par la méthode de Saltelli et al. [4] à partir des deux échantillons *A* et *B* ; pour chaque ligne *i* de l'échantillon, la valeur des paramètres C_1 à C_8 pour cette ligne permet de calculer une carte $Z^{(i)}$ à l'aide de la fonction `Campbell2D`.

Le nombre d'évaluations de la fonction test est :

$$N = 2 \cdot (3 + 1) \cdot M = 8 \cdot M$$

Le nombre d'appels à la fonction `Campbell2D` est *N*.

2.3 ■ Méthode *S*

La méthode *S* est celle proposée par Crosetto et al. [2]. Elle repose sur l'introduction d'une «variable interrupteur» ou variable «switch», que l'on note η , qui contrôle l'introduction d'une part d'incertitude dans le champ *Z*. On suit la procédure suivante :

1. on tire ensuite deux échantillons *A* et *B* de même taille *M* mêlant X_1, X_2 et une variable aléatoire binaire η pouvant prendre de manière équiprobable les valeurs 0 ou 1.
2. on estime les indices de Sobol de X_1, X_2 et de η par la méthode de Saltelli et al. [4] à partir des deux échantillons *A* et *B* ; pour chaque ligne *i* de l'échantillon, on suit la démarche suivante :
 - si $\eta^{(i)} = 0$, alors la valeur donnée à la carte $Z^{(i)}$ pour l'estimation de la sortie $Y^{(i)}$ est une valeur nominale constante, égale au champ *Z* généré pour le vecteur de paramètres C_1 à C_8 suivant : (5, 3, 1, -1, 5, 3, 1, -1).
 - si $\eta^{(i)} = 1$, alors la valeur donnée à la carte $Z^{(i)}$ pour l'estimation de la sortie $Y^{(i)}$ est une valeur aléatoire, obtenue en tirant les paramètres C_1 à C_8 dans leur loi de probabilité et en générant la carte $Z^{(i)} = Z(C_1, \dots, C_8)$ à l'aide de la fonction `Campbell2D`.

Le nombre d'évaluations de la fonction test est :

$$N = 2 \cdot (3 + 1) \cdot M = 8 \cdot M$$

Le nombre d'appels à la fonction `Campbell2D` est de l'ordre de $N/2$ (la fonction n'est appelée que quand $\eta = 1$).

3 Résultats

Les codes et fichiers se trouvent dans les dossiers `Resultats/23NovY1`, `Resultats/24NovY2` et `Resultats/24NovY3`.

3.1 ■ Comparaison des quatre méthodes d'estimation

On commence par comparer les estimations des indices de Sobol obtenus sur chacune des fonction test par les différentes méthodes d'estimation pour une même taille M d'échantillon initial. On fixe $M = 2048$ et on estime les indices de sensibilité pour Y_1 , Y_2 et Y_3 selon les quatre approches décrites en section 2 : L (avec un nombre $n = 1000$ de réalisations aléatoires de la carte Z), RD , RDG et S . Les résultats sont données en Tab. 1. S_{X_1} , ST_{X_1} , S_{X_2} et ST_{X_2} désignent les indices de sensibilité de premier ordre et totaux des facteurs scalaires X_1 et X_2 . S_Z et ST_Z désignent les mesures de la sensibilité du champ Z ; les grandeurs désignées par S_Z et ST_Z dépendent de la méthode utilisée. On indique :

- pour la méthode L : les indices de sensibilité de premier ordre $S_{\tilde{Z}}$ et totaux $ST_{\tilde{Z}}$ de la v.a. discrète \tilde{Z} , obtenus pour un nombre $n = 1000$ de réalisations aléatoires de la carte Z
- pour la méthode RD : la somme $\sum_i S_{C_i}$ des indices de premier ordre des paramètres scalaires C_i et la somme $\sum_i ST_{C_i}$ de leurs indices totaux
- pour la méthode RDG : les indices de sensibilité de premier ordre et totaux du groupe de facteurs scalaires (C_1, \dots, C_8).
- pour la méthode S : les indices de sensibilité S_η et ST_η de la variable binaire η

Les figures 2 à 13 donnent les résultats obtenus sous formes de graphes pour chaque méthode d'estimation et chaque fonction test.

Fonction	Méthode	S_{X_1}	S_{X_2}	S_Z	ST_{X_1}	ST_{X_2}	ST_Z
Y_1	L	0.850290213	-0.002400451	0.086159071	0.85400421	0.06653359	0.14921352
	RD	0.831227669	-0.007323054	-0.02939551	0.82528089	0.07106520	0.7841338
	RDG	0.831227669	-0.007323054	0.091048482	0.8252809	0.0710652	0.1714519
	S	0.904196858	-0.006681872	0.005971272	0.98693707	0.09534858	0.09948201
Y_2	L	0.1618556	0.5101516	0.3642044	0.1698186	0.4657908	0.3632460
	RD	0.156105433	0.469264153	0.31304476	0.177981954	0.471711507	0.4054459
	RDG	0.1561054	0.4692642	0.3493322	0.1779820	0.4717115	0.3750211
	S	0.1171483	0.5574584	0.1131128	0.3544458	0.7701271	0.3271639
Y_3	L	0.3910610	0.1607171	0.2610566	0.5444017	0.2550799	0.4013758
	RD	0.4085411505	0.1546305456	0.1637237	0.544435089	0.245480021	0.4202739
	RDG	0.4085412	0.1546305	0.2550215	0.5444351	0.2454800	0.3775389
	S	0.3615024	0.1433148	0.1183948	0.7123740	0.4789457	0.4311372

TAB. 1 – Estimation des indices de sensibilité pour les trois fonctions test.

3.2 ■ Convergence de l'approche de Lilburne & Tarantola

En plus de comparer la méthode L aux autres approches, on veut s'assurer de sa convergence lorsque le nombre n de réalisations aléatoires pré-générées du champ Z augmente. On suit donc la procédure suivante pour chaque fonction test f_1 , f_2 et f_3 :

- pour n valant successivement 3, 10, 25, 50, 100, 500 et 1000 :
 1. générer un jeu de n réalisations aléatoires de la carte d'entrée Z
 2. estimer les indices de sensibilité S_{X_1} , S_{X_2} , $S_{\bar{Z}}$, ST_{X_1} , ST_{X_2} et $ST_{\bar{Z}}$ à partir de ce jeu de n réalisations selon la méthode exposée en 2.1, avec $M = 2048$
 3. réitérer les étapes 1 et 2 un nombre $r = 20$ de fois
 4. calculer pour chaque indice de sensibilité sa moyenne empirique μ et son écart-type empirique σ sur les r replicas

Les résultats sont donnés en Tab.2, et sous forme graphique dans les figures Fig.2, Fig.6 et Fig.10.

4 Discussion

4.1 ■ Comparaison des quatre méthodes d'estimation

On prend comme valeurs de référence pour les indices de sensibilité les estimations données par la méthode RDG : en effet, dans ces trois cas tests, la réduction du champ Z aux seuls facteurs scalaires C_1 à C_8 se fait sans perte d'information, et les indices de sensibilité du groupe (C_1, \dots, C_8) mesure exactement l'influence de la variabilité du champ Z sur la variance de Y . On compare donc les résultats obtenus avec chacune des méthodes L , RD et S par rapport aux indices de référence obtenus par la méthode RDG .

Méthode L

Les méthodes L et RDG donnent des résultats identiques à $5 \cdot 10^{-2}$, pour tous les indices de sensibilité, quelque soit la fonction test. Ces résultats constituent donc une validation empirique de l'approche de Lilburne & Tarantola.

Méthode RD

Les méthodes RD et RDG donnent des résultats identiques (à 10^{-2} près) pour les indices S_{X_1} , S_{X_2} , ST_{X_1} et ST_{X_2} , quelque soit la fonction test. En revanche, la méthode RD donne des indices $S_{\bar{Z}}$ et $ST_{\bar{Z}}$ différents de ceux obtenus par la méthode RDG . En effet, pour la méthode RD , on a considéré la somme de S_{C_1} à S_{C_8} pour estimer l'indice $S_{\bar{Z}}$: ce faisant, on ne prend pas en compte l'ensemble des interactions entre les variables C_i , qui seraient capturées par les indices de second ordre S_{C_i, C_j} et d'ordres supérieurs. La somme des indices de premier ordre S_{C_1} à S_{C_8} est donc nécessairement inférieure à l'indice de premier ordre $S_{\bar{Z}}$. De même, on a considéré la somme des indices totaux ST_{C_1} à ST_{C_8} pour approcher l'indice total $ST_{\bar{Z}}$. Mais la somme des ST_{C_i} «compte plusieurs fois» les interactions entre les variables C_i , alors que l'indice $ST_{\bar{Z}}$ ne les compte qu'une seule fois. Ainsi l'indice total ST_{C_1} compte l'interaction entre C_1 et C_2 , de même que l'indice total ST_{C_2} : cette interaction est donc comptée deux fois dans la somme des ST_{C_i} . La somme des indices totaux ST_{C_1} à ST_{C_8} est donc nécessairement supérieure à l'indice total $ST_{\bar{Z}}$. La différence entre les méthodes RD et RDG est néanmoins moins marquée pour la fonction test Y_2 que pour la fonction test Y_1 .

Sortie	n	S_{X_1}	S_{X_2}	S_Z	ST_{X_1}	ST_{X_2}	ST_Z
Y_1	3	0.8866 ± 0.1180	0.0266 ± 0.0322	0.04694 ± 0.0564	0.8879419 ± 0.1188	0.06503683 ± 0.6449	0.08557542 ± 0.1005
	10	0.8375 ± 0.0888	0.0053 ± 0.0159	0.0859 ± 0.0488	0.8390 ± 0.0883	0.07504 ± 0.0404	0.1554 ± 0.0885
	25	0.8350543 ± 0.05302385	-0.0007420464 ± 0.005156668	0.09062465 ± 0.032722731	0.8371727 ± 0.05479156	0.07314243 ± 0.024013635	0.16485447 ± 0.05444106
	50	0.8366506 ± 0.03503150	0.0006850539 ± 0.007241824	0.09298079 ± 0.021936938	0.8351441 ± 0.03658607	0.07208710 ± 0.014994218	0.16269680 ± 0.03514308
	100	0.8230673 ± 0.02961339	-0.0002578500 ± 0.007525055	0.09770505 ± 0.018382354	0.8240857 ± 0.03115912	0.07810397 ± 0.013777128	0.17701135 ± 0.03189353
	500	0.8227824 ± 0.02050468	-0.0020904653 ± 0.011088110	0.10197176 ± 0.012730641	0.8237501 ± 0.02094132	0.07676039 ± 0.007598656	0.17941412 ± 0.01956842
	1000	0.8292831 ± 0.01730883	-0.0008786357 ± 0.010438493	0.09629757 ± 0.007636625	0.8309337 ± 0.01293915	0.07504477 ± 0.006167509	0.17095473 ± 0.01429712
Y_2	3	0.1970002 ± 0.07223544	0.5463787 ± 0.12307965	0.2363564 ± 0.14417332	0.2169972 ± 0.064336991	0.5466505 ± 0.122862787	0.2563747 ± 0.15685282
	10	0.1536693 ± 0.02883973	0.4787199 ± 0.06371981	0.3381120 ± 0.06792621	0.1829161 ± 0.026055594	0.4789663 ± 0.063125820	0.3671344 ± 0.07395985
	25	0.1595028 ± 0.03082381	0.4893578 ± 0.05405875	0.3230510 ± 0.06666291	0.1876929 ± 0.026464545	0.4891412 ± 0.055474748	0.3507544 ± 0.07173455
	50	0.1543404 ± 0.01702069	0.4849544 ± 0.02770119	0.3290898 ± 0.03132350	0.1833761 ± 0.016655715	0.4870611 ± 0.028155234	0.3586659 ± 0.03146773
	100	0.1545321 ± 0.01544147	0.4726348 ± 0.01992285	0.3406712 ± 0.02827587	0.1855625 ± 0.012467312	0.4732958 ± 0.020199703	0.3709438 ± 0.03058221
	500	0.1577492 ± 0.01460647	0.4735884 ± 0.01564200	0.3462898 ± 0.01197429	0.1842365 ± 0.004136197	0.4702980 ± 0.009762506	0.3729912 ± 0.01500561
	1000	0.1489888 ± 0.01415832	0.4677741 ± 0.01999	0.350901 ± 0.00735222	0.179523 ± 0.004955768	0.4689864 ± 0.008095741	0.3808107 ± 0.01172762
Y_3	3	0.4743984 ± 0.10960439	0.1896759 ± 0.04473472	0.1881723 ± 0.114864211	0.5958720 ± 0.085050845	0.2771428 ± 0.039056769	0.2768444 ± 0.16522280
	10	0.3961266 ± 0.05160446	0.1580559 ± 0.02239217	0.2645198 ± 0.051654044	0.5370375 ± 0.038590847	0.2515930 ± 0.019484664	0.3903721 ± 0.07593606
	25	0.4165438 ± 0.05322187	0.1593460 ± 0.02428625	0.2488137 ± 0.048971587	0.5528342 ± 0.037882219	0.2538291 ± 0.015236637	0.3678399 ± 0.07523875
	50	0.4162637 ± 0.02768506	0.1670069 ± 0.01487743	0.2539012 ± 0.038767089	0.5469098 ± 0.029252282	0.2561351 ± 0.014475771	0.3671836 ± 0.04620585
	100	0.4118655 ± 0.02812108	0.1650008 ± 0.01816890	0.2578474 ± 0.021302241	0.5444607 ± 0.016224587	0.2524560 ± 0.010782726	0.3729858 ± 0.03393773
	500	0.4109903 ± 0.02112225	0.1643341 ± 0.02778258	0.2580254 ± 0.012581969	0.5449512 ± 0.010428179	0.2558176 ± 0.007436659	0.3746230 ± 0.01794222
	1000	0.4015318 ± 0.01594067	0.1608713 ± 0.01176296	0.2655688 ± 0.007805696	0.5416655 ± 0.008268677	0.2512392 ± 0.005784579	0.3904592 ± 0.01239766

TAB. 2 – Estimation des indices de sensibilité par la méthode L pour un nombre n de réalisations aléatoires de la carte Z croissant. Pour chaque fonction test, chaque n et chaque indice, on donne la moyenne empirique sur 20 réplicas ± l'écart-type empirique.

Méthode S

Les estimations des indices de sensibilité par l'approche S sont significativement différentes des valeurs de référence. Cette méthode sur-estime systématiquement l'indice de premier ordre S_Z et sous-estime systématiquement

les indices totaux ST_{X_1} et ST_{X_2} , quelque soit la fonction test. L'approche S ne semble donc pas appropriée pour l'estimation des indices de Sobol dans un modèle avec variables d'entrée distribuées spatialement.

4.2 ■ Convergence des estimateurs de Lilburne avec n

La table Tab.2 et les figures Fig.2, Fig.6 et Fig.10 montrent la convergence des estimateurs des indices de Sobol S_{X_1} , S_{X_2} et S_Z , estimés par la méthode L , lorsque le nombre n de réalisations aléatoires du champ spatial Z augmente. Les résultats montrent deux choses :

1. la précision des estimateurs des indices augmente avec n ; sur les trois fonctions test f_1 , f_2 et f_3 , l'écart-type sur les $r = 20$ répliques est inférieur à 0.05 pour $n \geq 50$, quelque soit l'indice considéré (S_{X_1} , S_{X_2} , S_Z , ST_{X_1} , ST_{X_2} ou ST_Z).
2. pour un nombre n de cartes faible ($n < 50$), les estimateurs sont biaisés : tendance à la sous-estimation de S_Z et ST_Z , et à la sur-estimation de S_{X_1} , S_{X_2} , ST_{X_1} et ST_{X_2} , en prenant pour référence les indices de sensibilité obtenus par la méthode RDG . Pour n plus grand, ce biais diminue, et semble tendre vers 0.

5 Conclusion

Conclusion : D'après ces résultats, on peut conclure que

- l'approche de Lilburne & Tarantola (méthode L) est valide, en ce sens que pour n suffisamment grand ($n \geq 100$), elle reproduit les indices de sensibilité obtenus par la méthode RDG (à 10^{-2} près)
- sur les trois fonctions test retenues, le nombre de cartes suffisant est $n = 100$ pour obtenir un écart-type inférieur à 0.05 sur la valeur des indices de sensibilité (pour différents jeux de n réalisations aléatoires de la carte Z)
- pour n petit, l'approche de Lilburne & Tarantola sous-estime l'influence du champ Z et les estimateurs sont peu précis (fort écart-type)
- la méthode RD ne permet pas de discuter de manière satisfaisante de l'influence du champ Z sur la variabilité de Y
- la méthode S doit être rejetée

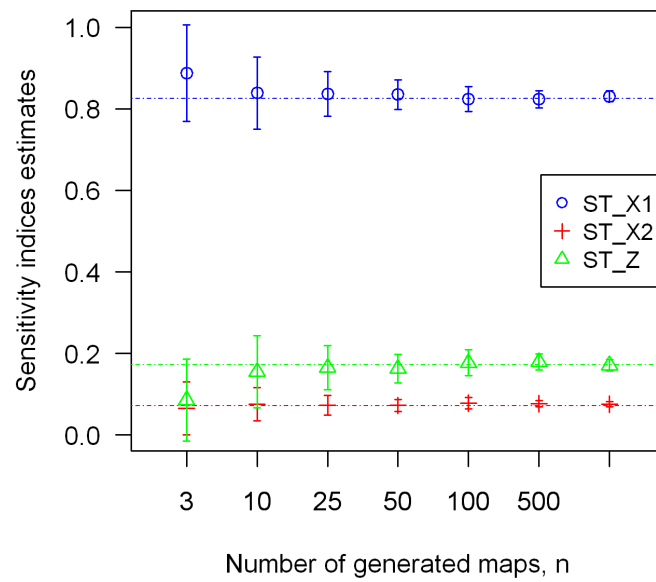
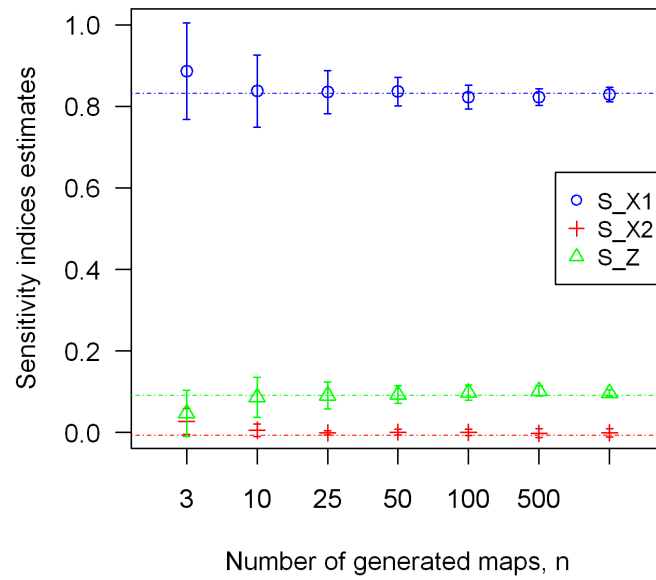


FIG. 2 – Convergence des estimateurs des indices pour la méthode L
 Fonction test Y_1 - Les lignes pointillées indiquent l'estimation des indices par la méthode RDG .

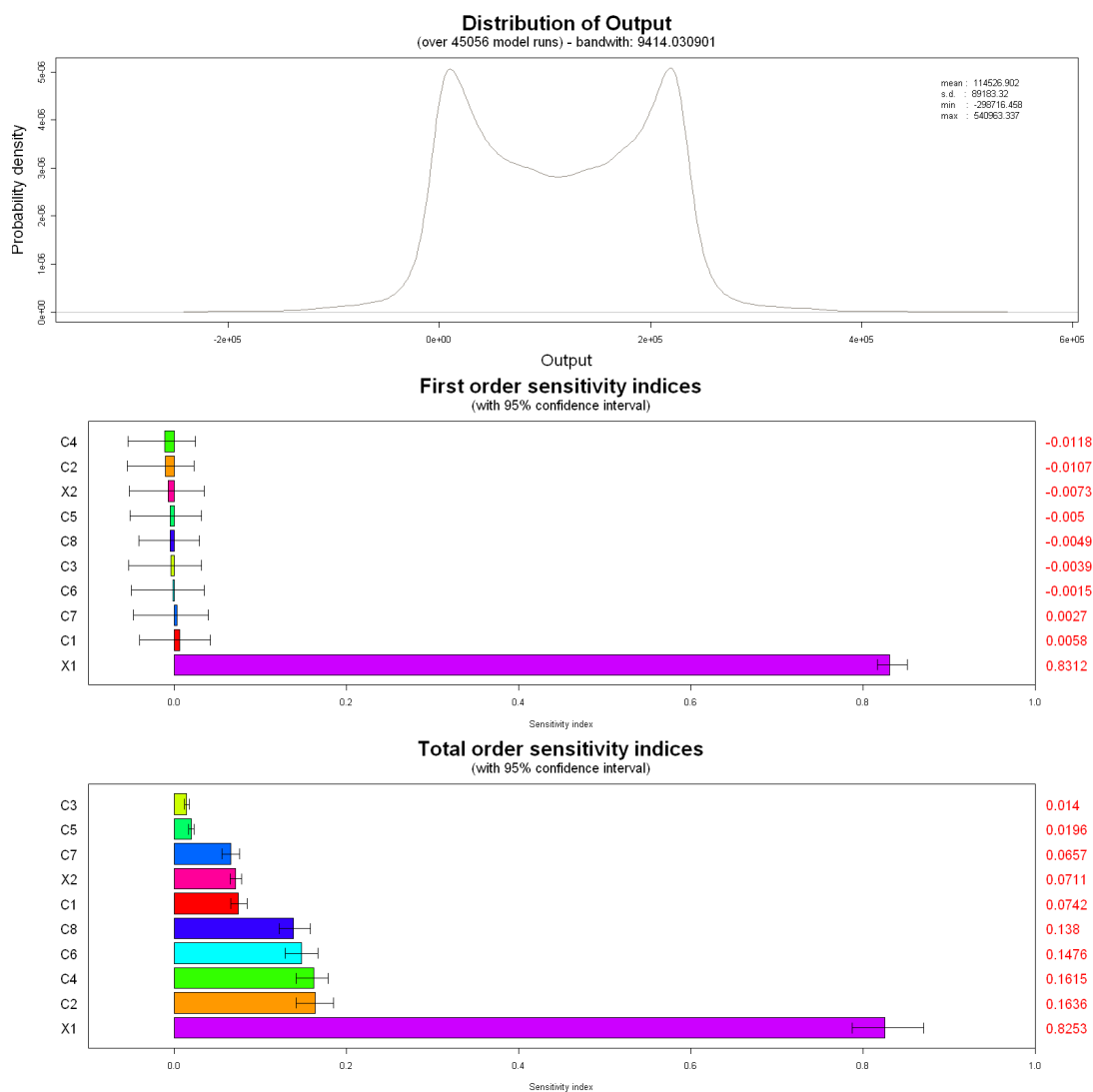


FIG. 3 – Estimateurs des indices pour la méthode *RD*
Fonction test Y_1

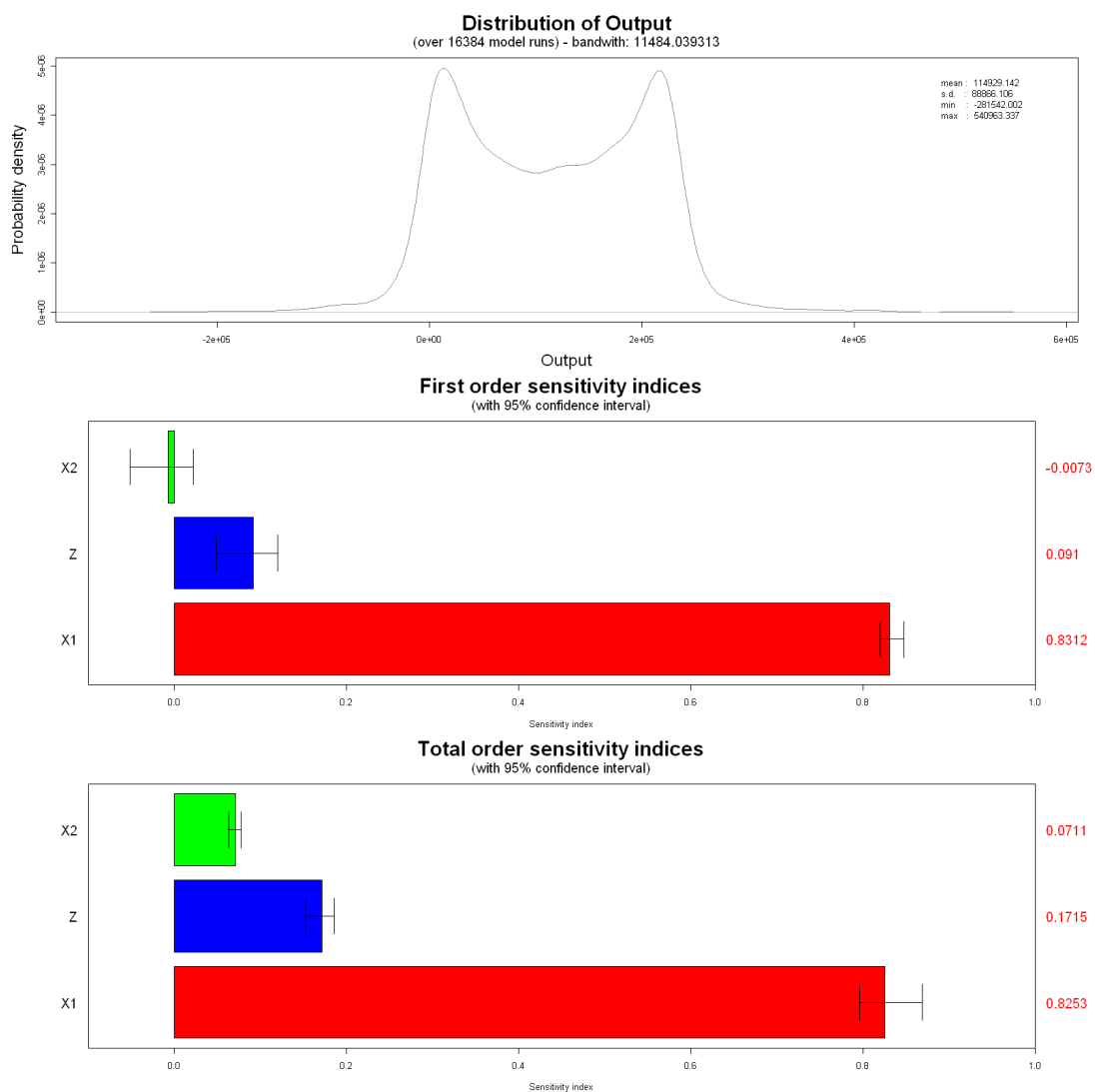


FIG. 4 – Estimateurs des indices pour la méthode *RDG*
Fonction test Y_1

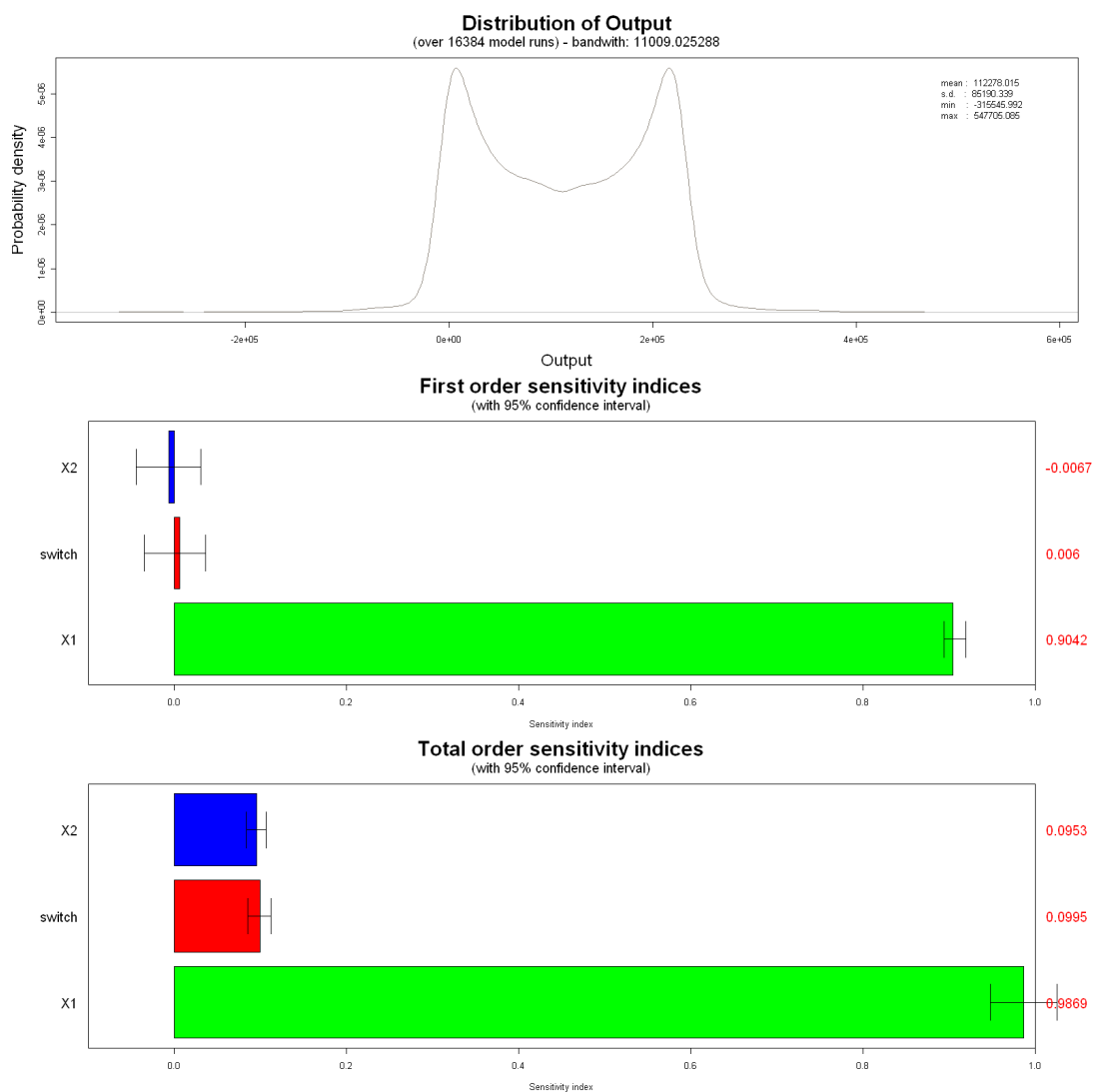


FIG. 5 – Estimateurs des indices pour la méthode S
Fonction test Y_1

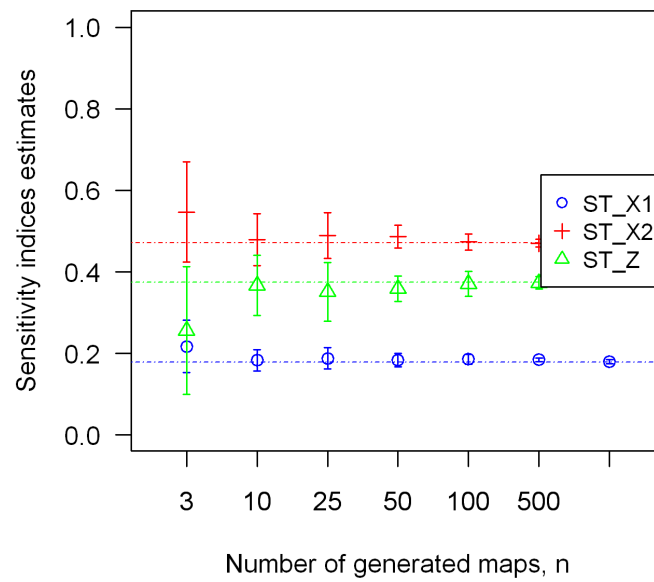
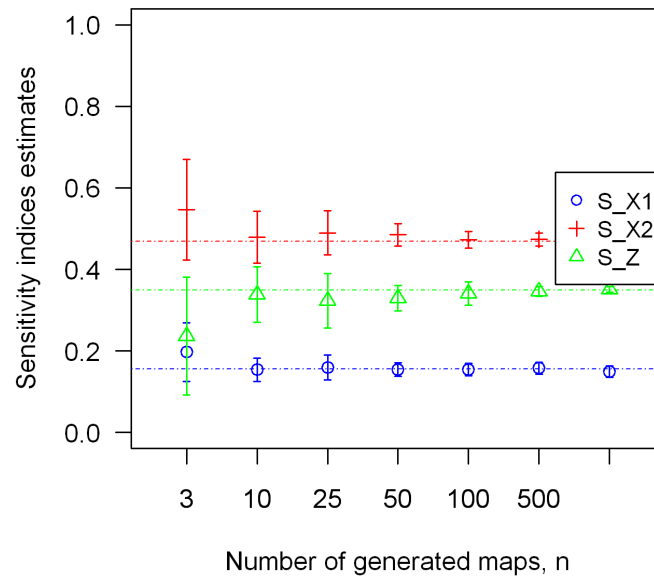


FIG. 6 – Convergence des estimateurs des indices pour la méthode L
Fonction test Y_2 - Les lignes pointillées indiquent l'estimation des indices par la méthode *RDG*.

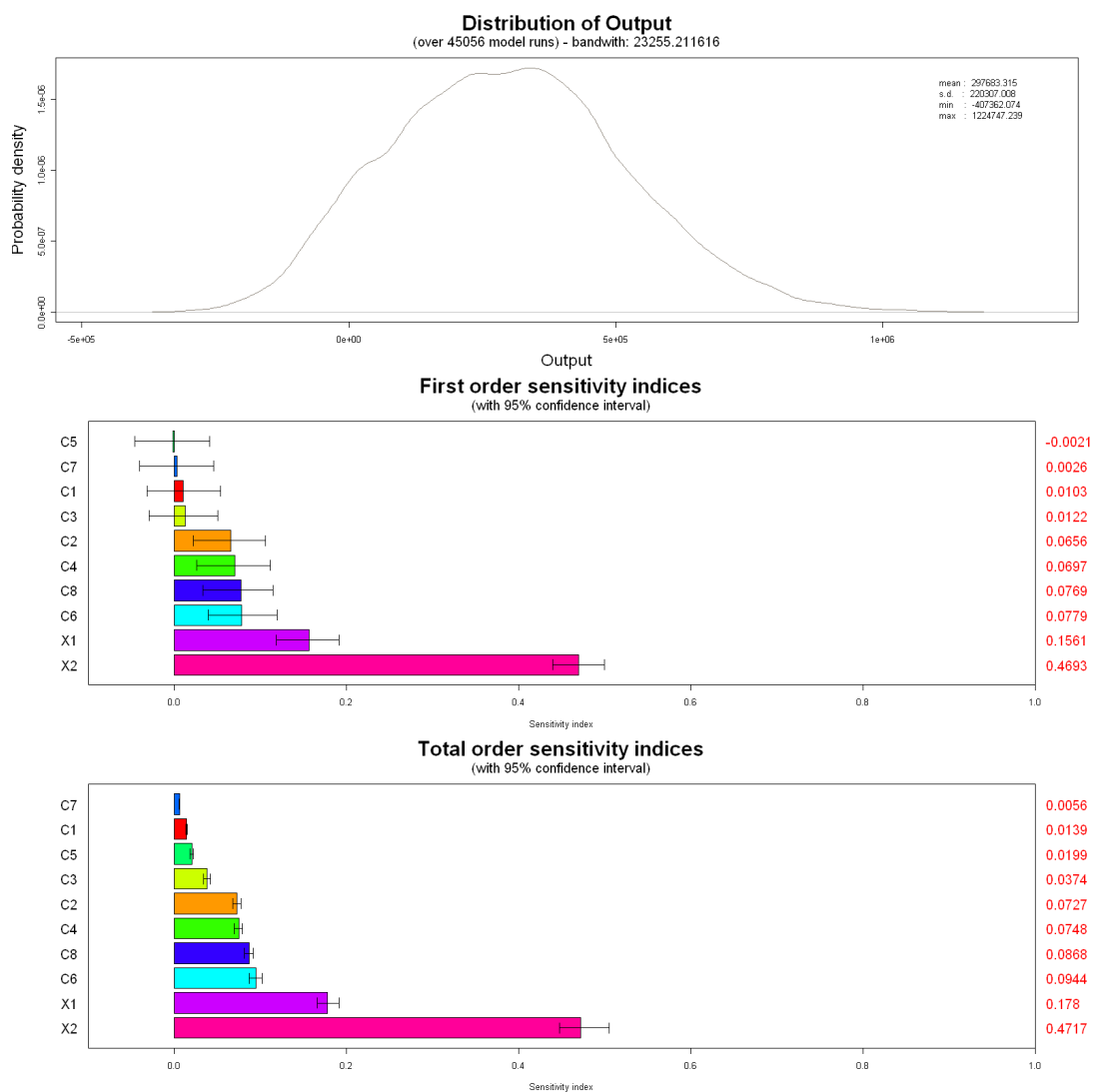


FIG. 7 – Estimateurs des indices pour la méthode MP
Fonction test Y_2

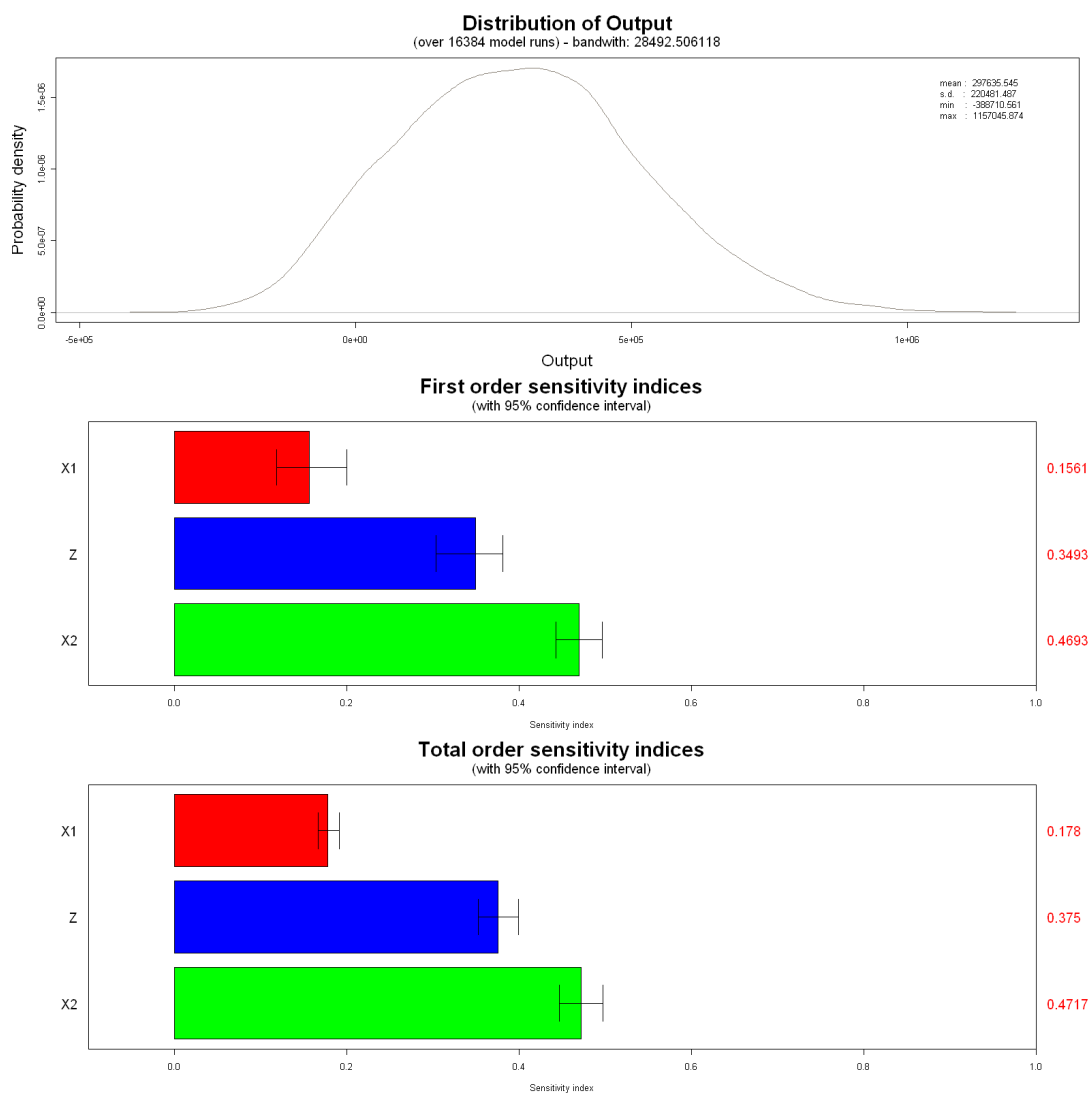


FIG. 8 – Estimateurs des indices pour la méthode MPG
Fonction test Y_2

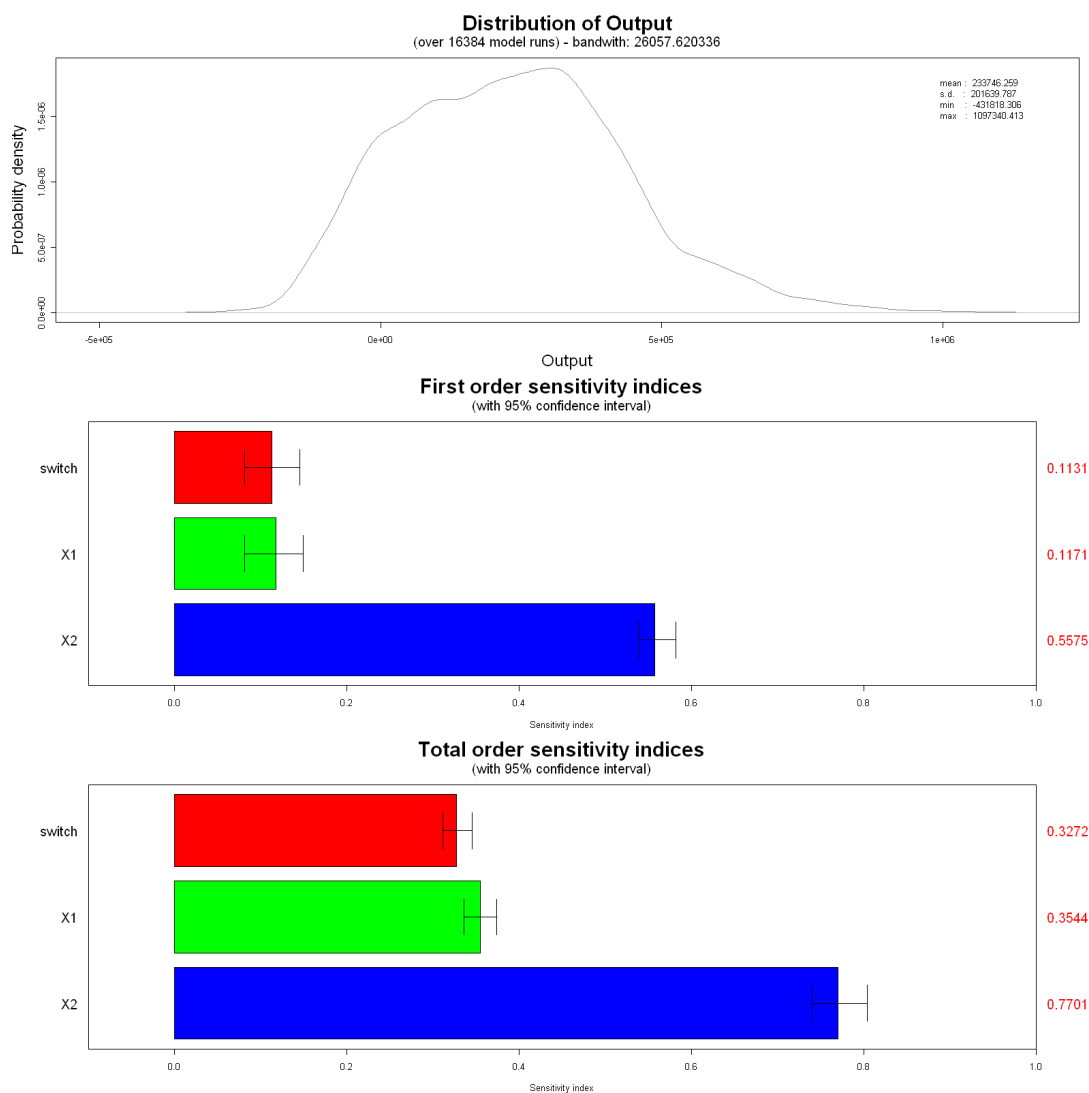


FIG. 9 – Estimateurs des indices pour la méthode S
Fonction test Y_2

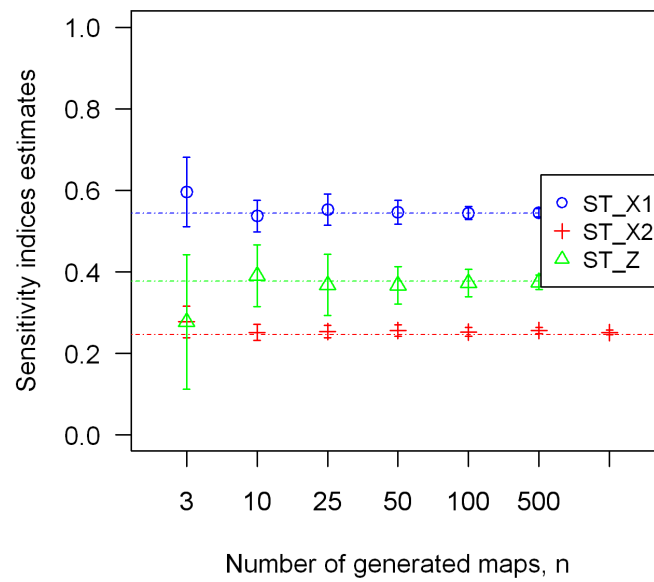
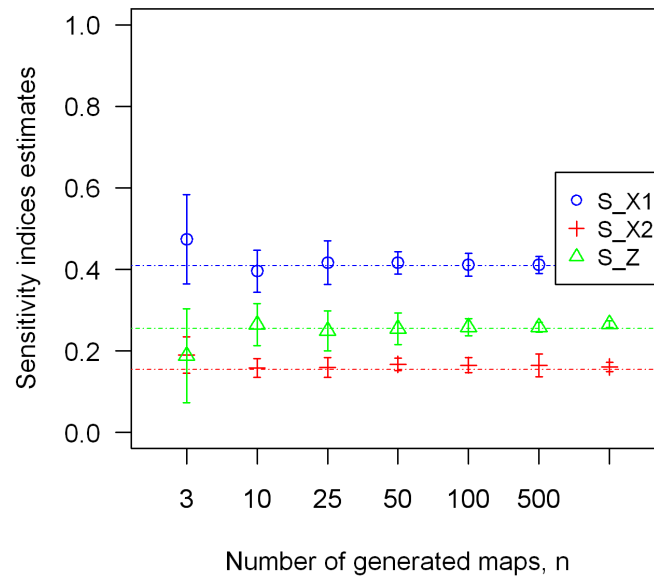


FIG. 10 – Convergence des estimateurs des indices pour la méthode L
Fonction test Y_3 - Les lignes pointillées indiquent l'estimation des indices par la méthode *RDG*.

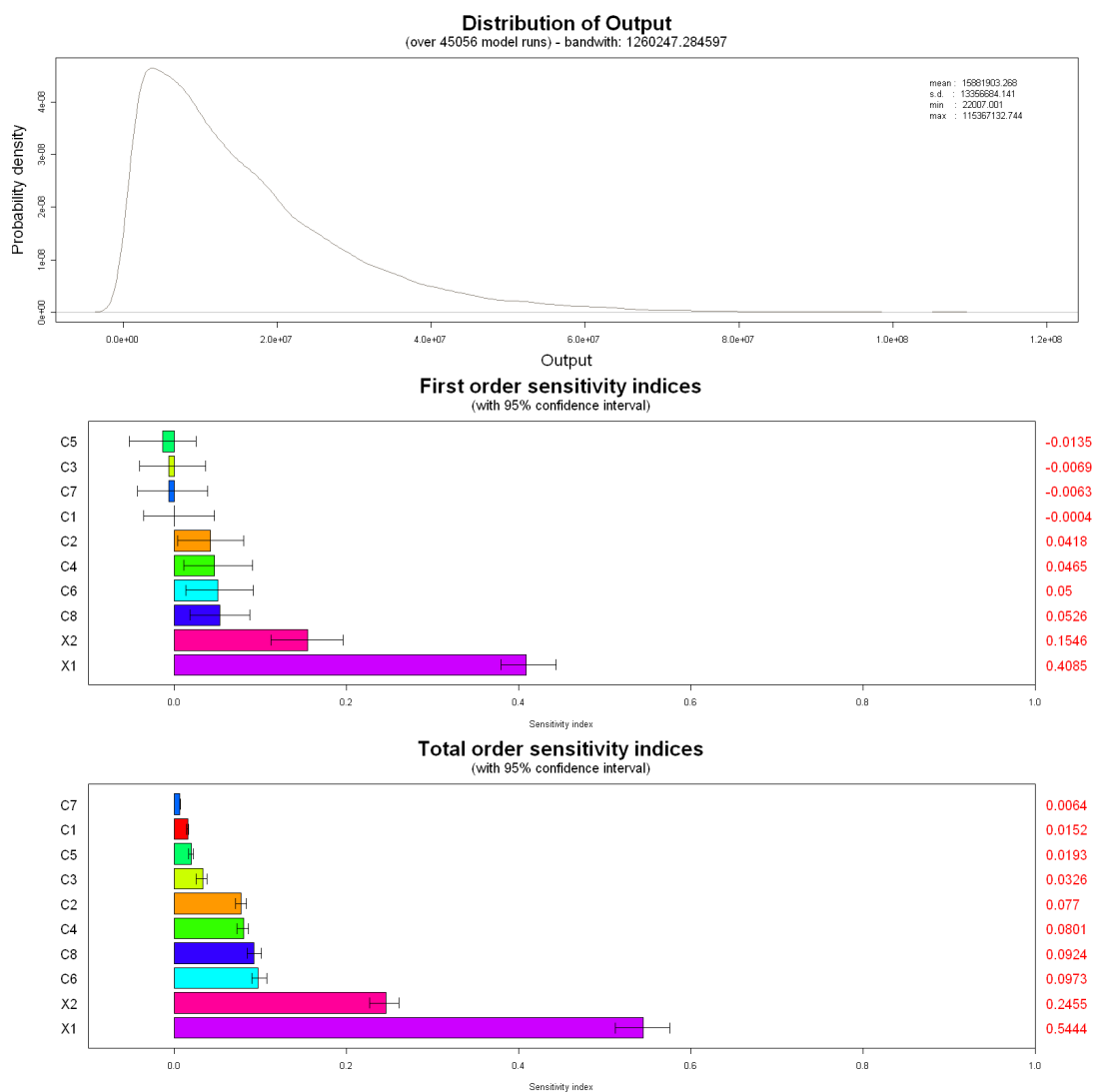


FIG. 11 – Estimateurs des indices pour la méthode MP
Fonction test Y_3

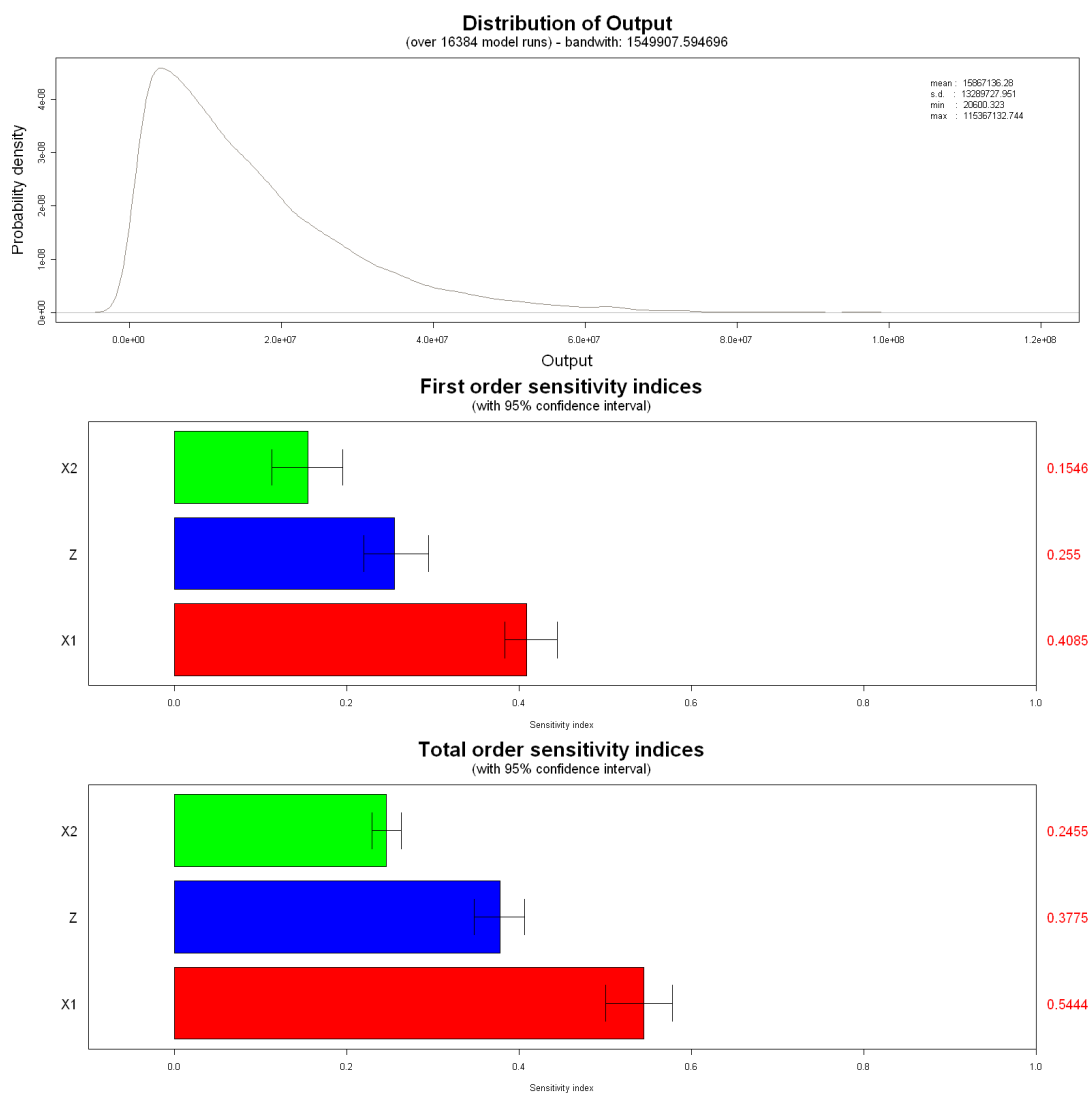


FIG. 12 – Estimateurs des indices pour la méthode MPG
Fonction test Y_3

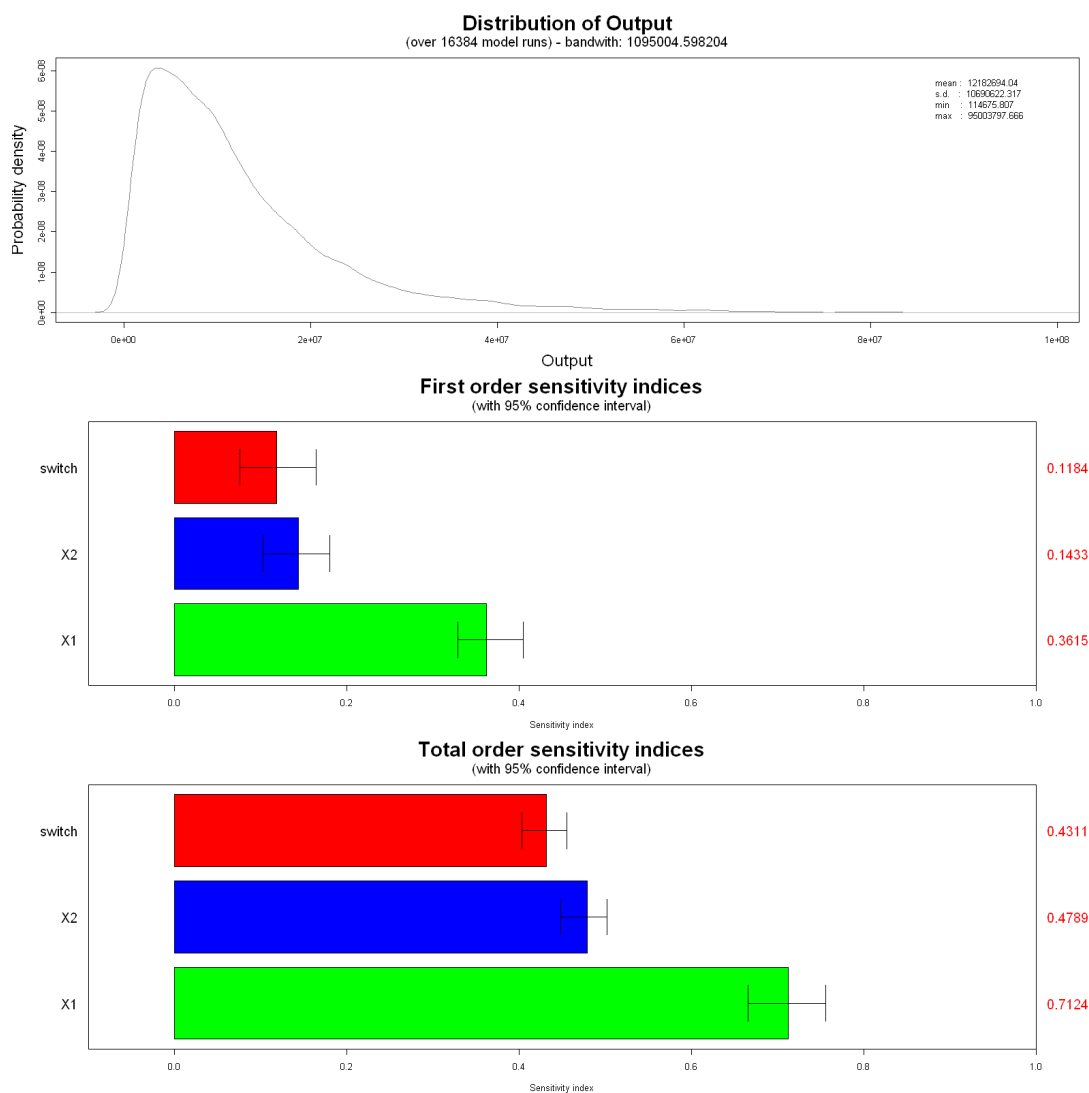


FIG. 13 – Estimateurs des indices pour la méthode S
Fonction test Y_3

Bibliographie

- [1] D. Busby, T. Romary, M. Feraille, and S. Touzani. An integrated approach for uncertainty and sensitivity analysis in reservoir forecasting. *Soumis à Journal of Comput. Geosci.*
- [2] M. Crosetto and S. Tarantola. Uncertainty and sensitivity analysis : tools for GIS-based model implementation. *Int. J. Geographical Information Science*, 15 :415–437, 2001.
- [3] L. Lilburne and S. Tarantola. Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, 23 :2 :151–168, 2009.
- [4] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis - The Primer*. Wiley, 2008.
- [5] E. Volkova, F. Van Dorpe, and B. Iooss. Modélisation du transport de sr90 en milieu poreux saturé et analyse de sensibilité du modèle : application sur un site de stockage temporaire de déchets radioactifs (crr kurchatov institute, russie). Technical report, CEA, 2005.

A	Code	23
----------	-----------------------	-----------

A Code

Le code se structure de la manière suivante :

- `main.r` : définition des paramètres généraux de l'étude et contrôle de l'exécution des autres scripts
- `toyfunctions.r` : définition de la fonction $Z(\theta, \phi)$, des fonctions tests et des lois de probabilité des variables d'entrée
- `analyses.r` : estimation des indices de Sobol pour chacune des quatre méthodes
- `SortiesGraphiquesLilburne.r` : synthèse des résultats pour la méthode Lilburne et création d'un graphique présentant ces résultats

